Relational contracts when the agent's productivity inside the relationship is correlated with outside opportunities

Alexander F. Wagner*

September 15, 2013

Abstract

An agent can choose to forego benefits from side opportunities and to instead provide benefits to the principal. In return, the principal offers rewards. If this exchange is not contractible, typically repeated interaction will be required to sustain it. This model allows the agent's productivity in contractible and possibly also non-contractible actions inside the relationship to be correlated with productivity in side activities. This arguably realistic assumption yields several novel implications for the feasibility of relational contracts and for agent selection by principals. The analysis reveals, for example, that optimal agent productivity is often non-monotonic in the importance, to the principal, of ensuring agent reliability.

JEL-Classification: C73, D82, M51.

Keywords: Relational contracts, agent selection

*Assistant Professor of Finance at the University of Zurich, Research Affiliate at CEPR, and Research Associate at ECGI and at the Harvard Kennedy School. Office address: Department of Banking and Finance, Swiss Finance Institute - University of Zurich, Plattenstrasse 14, CH-8032 Zurich, Switzerland. Phone: +41-44-634-3963, Email: alexander.wagner@bf.uzh.ch. This paper was previously circulated under the title "Loyalty and Competence." I thank Philippe Aghion for suggesting to consider the tradeoff between loyalty and competence, and Alberto Alesina, Nolan Miller and in particular Richard Zeckhauser for encouragement and comments. Mario Häfeli provided excellent research assistance. Helpful comments were also provided by Randall Akee, George Baker, Charles Cohen, Robert Gibbons, Marina Halac, Oliver Hart, Sonal Pandya, Catherine Thomas, Robert Urbatsch, Lucy White, and seminar participants at Boston University, George Mason University, Harvard, MIT, the University of Chicago, and the University of Zurich, as well as participants at several conferences. This research was supported by the Austrian Academy of Sciences (APART fellowship), the NCCR FINRISK, UZH URPP "Finance and Financial Markets," and the Swiss Finance Institute.

1 Introduction

Relations are the lifeblood of organizations. Cooperative behavior can often only be sustained in repeated relationships (Axelrod 1984). Relations can also foster attachment to organizations, thus facilitating the voice option rather than the exit option for the members of the organization which can be particularly important in times of crisis (Hirschman 1971). More generally, relations often help achieve outcomes that are otherwise not available due to high transaction costs.

One way in which economists have successfully modeled relations is through self-enforcing (relational) contracts. Seminal contributions have established conditions for the feasibility of stationary relational contracts (Bull 1987, MacLeod and Malcomson 1989, Levin 2003). Models of self-enforcing contracts have been fruitfully employed to analyze relations both within organizations (e.g., Rayo (2007)) and between (e.g., Doornik (2006)), and to study the implications of the interaction between formal contracts and informal agreements for firm boundaries (Baker, Gibbons, and Murphy 1994, Baker, Gibbons, and Murphy 2002). Recent important extensions of the basic theory have allowed for privately observed effort and output (MacLeod 2003, Fuchs 2007), privately observed costs of making promised payments (Englmaier and Segal 2012, Li and Matouschek 2012), unobservable value of the relationship (Halac 2012), limited liability (Fong and Li 2012), as well as non-stationarity of contracts (Yang 2009, Thomas and Worrall 2010) and the building of relationships (Chassang 2010).

This paper starts from the traditional framework of stationary contracts with observable effort and output, but introduces a feature of many real-world situations that has hitherto not been explicitly analyzed. In particular, economic models of relational contracts typically abstract away from the possibility that an agent's productivity in tasks inside the relationship – both in those tasks that are the focus of the self-enforcing contract and in those that are contractible – may be correlated with his side opportunities. This paper allows for this possibility and shows that it can have important implications for agent selection and for relations.

To develop the analysis, consider an organization, for example, a consultancy, where a (female) manager (the principal, M) and a (male) worker (the agent, W) produce output. M chooses W's *competence*, that is, M selects W. The more *competent* W is, the more contractible output he produces; he writes better reports for the clients of the consultancy. W's competence is a characteristic (not W's action). W regularly needs to decide whether to bear personal (opportunity) costs for the principal. To fix terminology, let us say that W acts *reliably* when he is willing to bear personal costs to benefit M. (It is convenient to have a single short-cut for referring to this behavior throughout this paper. Other terms could be used, depending on the specific application this model is adapted to. Only in some settings is "effort" a natural interpretation of a reliable action.) Thus, reliability is an optimally chosen behavior, not a given trait. W acts *independently* when he instead chooses to obtain private benefits, perhaps by voting for his preferred option or by not participating in the vote and pursuing matters of private interest or by using a meeting with a client to establish some personal business relationships. More competent agents have more such opportunities for private benefits, that is, typically there is a positive correlation between inside and outside productivity. The extent to which W's competence affects the value for M of W's reliability depends on the context. In some cases, such as when reliability means voting for M's proposal in a group meeting or lauding M's abilities in a client meeting, W's competence is irrelevant for the value that M derives from W's reliability. In other cases, a more competent W will produce greater reliability benefits for M than a less competent W. W is willing to behave reliably and forego his side opportunities only if he receives appropriate rewards from M. Rarely will M and W be able to write a binding contract about all aspects of allegiance and its rewards. Rather, important parts of reliability and its results are usually non-contractible, requiring repeated interaction to sustain it.

The analysis shows first that the feasibility of reliability tends to be limited if a highly competent W requires too costly rewards for reliability (due to his better outside options on the reneging path) but if the value W creates for M through his reliability does not vary much with his competence. In this case, M may have too strong an incentive to renege on promised reliability rewards. Thus, a tradeoff between competence and reliability arises.The model implies that a more competent W has endogenously *higher* costs of reliable actions in this case (in contrast to a standard efficiency wage setting where cost of effort is decreasing in competence). If, by contrast, W's competence increases the value of his reliability to M sufficiently strongly, this creates enough opportunities for M to reward W for reliability. In sum, if the surplus defined by the spread between the value of reliability to M net of the cost of reliability to W, adjusted for discounting, falls too fast with W's competence, a relational contract cannot be sustained even with a highly competent W.

The model next yields empirical predictions for the optimal level of competence from M's perspective. Existing work on relational contracts does of course recognize that higher outside options make relations more difficult. But because the existing literature on self-enforcement does not allow for correlations between inside and outside productivity, taken literally, these models make the counter-factual prediction that it is optimal to pick the counterparty with the lowest possible reservation utility as the partner for interaction.

In the present framework, the empirical predictions instead naturally depend on the way competence affects the value of reliability. The most interesting comparative statics arise for the case where a tradeoff between reliability and competence exists. Competence is always (weakly) lower under reliability than under independence. On the one hand, M may willingly sacrifice some worker competence in order to secure reliable behavior. On the other hand, reliability need not be optimal because it may require M to give up too much worker competence. Thus, even where reliability is feasible and has net positive value for the manager, the manager may prefer a more competent worker with whom she keeps repeating the statically optimal equilibrium of the reliability game. In particular, independence and, thus, maximal competence is optimal when reliability is not so important to the principal.

Notably, however, in the range where reliability is optimal, when M values reliability more, she will actually *increase* W's competence. This surprising result occurs because with a higher valuation of reliability M can more credibly promise rewards. Overall, therefore, there arises a non-monotonic relationship between the value of reliability to the principal and optimal agent competence. While other model classes presumably can also generate such non-monotonicities under some conditions, the contribution here is to show that where a relational contracts model is seen as appropriate to study a given setting, the correlation of inside and outside productivity can yield an outcome that does not occur in the absence of this correlation. In particular, this finding can be compared with the prediction available, by analogy, from other work that contains a version of the notion that access to outside options affects the feasibility of relational contracts. Specifically, Baker, Gibbons, and Murphy (2002) show that asset ownership has an effect on the feasibility of relational contracts between firms. To see the relation between and the different implications of Baker, Gibbons, and Murphy (2002) and the present paper, note that non-integration between an upstream (supplier) and downstream (buyer) firm enhances the upstream party's "competence" in the sense that it can work on its outside options more effectively than it can under integration. In Baker, Gibbons, and Murphy (2002), where reliable delivery of the product from the upstream party to the downstream party is important, integration is more likely. Because that framework naturally considers only two levels of "competence" (integration and non-integration), it suggests a monotonic relationship between the value of ensuring reliable delivery and competence in terms of asset ownership. The present paper shows that this intuition does not extend to the case of continuous degrees of competence.

Finally, this paper also shows that independence and reliability can be signals of competence (to an uninformed outside market) made effective by the different marginal costs of reliability that agents of differing competence levels face.

Besides contributing to the literature on relational contracts, the paper also is related to some other work. First, the general idea that participation constraints can be type-dependent was analyzed by Jullien (2000) for the case of explicit contracts; here, relational (implicit) contracts are the focus. Second, many papers assume that agents differ in their preferences for taking into consideration the preferences of the partner (for example, Aghion and Tirole (1997) and Fehr and Falk (2002)). Here, by contrast, the costs of acting to the benefit of others is endogenously determined by competence. Third, some studies consider the case where principals care about both economic and social competence, broadly defined, of their counterparts, but focus on different issues than the present one. For example, reliable employees are "yes-men," but in contrast to Prendergast (1993), the present model does not entail an informational story. Firm-specific capital (Grossman and Hart 1986, Hart 1995) need not make a worker more reliable; when internal opportunities for applying the high firm-specific knowledge arise, such a worker may be quite disloyal to his boss, by, for example, taking a job in another division. Austen-Smith and Fryer (2005) emphasize how investment in human capital can make an agent unattractive for a social group. There are two main novel aspects of the present analysis. First, even though competence matters also to the production for M, M may willingly forego worker competence. Second, reliability here is a matter of choice that is endogenously determined by competence.

Finally, other model frameworks that do not use relational contracts also yield the prediction that principals may hire less than fully competent workers. For example, Prendergast and Topel (1996) show that a principal who values the power to affect his agent's welfare may prefer less competent agents; Glazer (2002) presents a static model of rent-division that focuses on the optimal mix of internal and external rent-seeking abilities; Friebel and Raith (2004) argue that a manager may fear competent subordinates because they may wish to take his post; and Egorov and Sonin (2011) use a static model with asymmetric information to demonstrate that the higher the stakes are for dictators, the more they may favor incompetent viziers who are more "loyal" (a term that could be used to describe the agent's behavior also in the present model). These models do not allow the worker to choose to be reliable, they do not allow for the manager's attempts to foster reliability, and they have different implications and empirical predictions. In the political economy context, some evidence (see Section 4) exists that is in line with the relational contracts framework's unique prediction that, where reliability is optimal, when M needs more reliability, she will in fact be able to and prefer to hire more competent workers.

The paper is organized as follows. Section 2 sets up the model and Section 3 analyzes it. Sections 4 and 5 discuss applications. Section 6 offers concluding remarks.

2 Model setup

The model considers the interaction between a manager (M) and a worker (W). Both the manager and the worker are risk-neutral and infinitely-lived with common discount factor

$\delta \in (0,1)$. Risk aversion is not necessary to create a potentially beneficial role for the relation, and so it is excluded. The worker's characteristic is his competence level, $\theta \in [0,1]$. The manager chooses the worker's competence. A natural interpretation of this assumption is that M hires exactly one W. W can take exactly one job. We can also think of M as choosing W's training, which is assumed to have a fixed cost. Over the course of the game, W's competence stays fixed. M has all the bargaining power.

Basic payoffs. M is the residual claimant of output. In each period, W produces output $\bar{y}$ with probability θ and $\underline{y}$ with probability $(1-\theta)$, where $\Delta y = \bar{y} - \underline{y} > 0$.[1] This output is contractible. As a model short-cut, the output flows solely from competence. M pays a spot wage w to W, as discussed below. Thus, for a given θ, M's expected payoff before calculating benefits and costs of reliability is $E_\theta y - w = \underline{y} + \theta \Delta y - w$. Similarly, W's payoff in the absence of reliability is w.

M's gains from W's reliability. The manager can, in each period, realize an additional potential gain (or avoided loss) $v(\theta) > 0$, where v is commonly known and commonly observed when it accrues to M. Assume that $v'(\theta) \geq 0$.

However, allowing M to obtain $v(\theta)$ requires W to bear some personal costs. These costs of cooperation are here modeled as foregoing the side opportunities $b(\theta)$.

W's side opportunities. In each period, a side opportunity of value $b(\theta) > 0$ arises. Assume $b'(\theta) > 0$. That is, more competent individuals have better and more side opportunities. If W chooses to pursue this opportunity, he acts *independently.* If W does not use the side opportunity, he acts *reliably.* Throughout, everybody can observe whether the worker is reliable. The fact that W has to bear costs in order for M to realize the gains from reliability is a defining characteristic of reliability. Reliability has *positive social (joint) value* when $v(\theta) > b(\theta)$.

The side opportunities can take many forms. For example, a highly competent employee

[1] Alternatively, this assumption can be interpreted as saying that W produces, with a fixed amount of time (normalized to 1), θ units of high-quality output and $(1-\theta)$ of low-quality output.

of a software company may engage in consulting activities; an employee of a regulatory agency can give a presentation on his field of expertise at a conference of professionals.

The wage w and expected private benefits $b(\theta)$ are additively separable in the worker's utility function. It does not directly harm M when W uses his side opportunity.

The model uses a simple, reduced form of the various roles of competence. A richer framework would endogenize $v(\theta)$ and $b(\theta)$, though this would come at the cost of additional complexity.

Is reliability a kind of effort? One might think of W's decision as his effort choice. This is a valid interpretation as W and M are in an agency relationship. However, the equilibrium costs of effort for W will be determined by M's choice of W's competence level and the ensuing strategic interaction with M. This will ultimately yield different predictions for how the cost of effort depends on competence than is implicit in a standard efficiency-wage model. In the light of these different results, a different term than "effort" seems appropriate.

Rewards for reliability. The manager needs to offer the worker something in exchange for bearing personal costs. The most direct way to accomplish this is for the manager to pay the worker for reliability, and this is the case the analysis focuses on. The monetized value of this reward – say that of a company car, or invitations to basketball games – is given by $x \geq 0$, and is a choice variable for M. It is assumed to be stationary for reasons discussed later. It is also commonly observed.

W's reservation utility. W's reservation utility is assumed to be:

$$\bar{u}(\theta) = \theta\bar{y} + (1-\theta)\underline{y} + \theta\alpha. \qquad (1)$$

The first part of this formula, $\theta\bar{y} + (1-\theta)\underline{y}$, is W's expected productivity. One way to rationalize the assumption that in alternative employment W receives this amount is to posit that W's alternative job opportunity is self-employment, where neither reliability nor side opportunities are relevant.[2] The assumption that $b(\theta)$ does not play a role in W's reservation utility will imply that M can economize on base wages. It is, in essence, a short-cut to obtain the (intuitive) baseline result that there are benefits for M to get a high-competence worker if he is affordable.

In the second part of equation (1), the parameter $\alpha < 0$ measures the marginal product of W's competence in this alternative employment *relative to when he works for M*; thus, α says something about the quality of the match between M and W. The analysis here focuses on the case where $\alpha < 0$. A negative α indicates that M has a relative advantage at using W's competence in the contractible part of production, that is M is *relatively productive*. This is the interesting case; after all, if M were less productive than the market, an obvious reason to hire a less competent W is that highly competent W's are too expensive because others are willing to pay them more. Note that the assumption $\alpha < 0$ makes no statement about the marginal impact of competence on side opportunities or on the value of reliability. (Results for $\alpha > 0$ are available on request.)

A break-up of the relationship is in principle costless. However, M's reservation utility is assumed to be sufficiently low to make the M-W match attractive for M, and it is always in M's interest to hire W, if just for "spot interaction" (defined below).

Summary. Each period, M and W face a game characterized by the stage game payoffs given in Table 1, where payoffs in each cell are in the format U_M, U_W. (The timing of moves is discussed below.)

Table 1: Payoff matrix in the reliability game

	W acts reliably	W acts independently
M rewards	$E_\theta y + v(\theta) - [w + x]$, $w + x$	$E_\theta y - [w + x]$, $w + b(\theta) + x$
M does not reward	$E_\theta y + v(\theta) - [w]$, w	$E_\theta y - [w]$, $w + b(\theta)$

[2] It is beyond the scope of this paper to develop a full competitive labor market equilibrium model in which M may also enter relational contracts in other firms.

3 Analysis

3.1 Statically optimal behavior (spot interaction)

For future reference, define *spot interaction* as the equilibrium that arises out of statically optimal behavior by the players. When M and W cannot contract on reliability, this would be the only equilibrium that arises if indeed M and W were meeting for only one round.[3] In this equilibrium, M does not pay for reliability ($x = 0$), and it is thus optimal for W to act independently, which in turn rationalizes M's strategy. Anticipating this equilibrium, M chooses the optimal competence level by solving

$$\max_{\theta} U_M = \underline{y} + \theta \Delta y - w \tag{2}$$
$$s.t.\ U_W = w + b(\theta) \geq \bar{u}(\theta)$$

The constraint states that inside utility, given by wages plus the expected value of side opportunities, must, on average, be no less than reservation utility. To minimize wage costs, M will offer a spot wage

$$w = \bar{u}(\theta) - b(\theta), \tag{3}$$

i.e., she will make W's participation constraint bind, where she can benefit from the fact that highly competent workers can engage in more side opportunities. Simplifying (2) by using the definition of $\bar{u}(\theta) = \theta\bar{y} + (1-\theta)\underline{y} + \theta\alpha$, M's problem is equivalent to

$$\max_{\theta} b(\theta) - \theta\alpha. \tag{4}$$

Because $\alpha < 0$, $\theta = 1$ is optimal, leading to a payoff of $b(1) - \alpha$.

[3] The reference to this equilibrium as having "spot" characteristics is meant to distinguish it from a relational equilibrium which requires long-term interactions. The literature sometimes uses the terminology of "no trade" (MacLeod and Malcomson 1998) for the notion of statically optimal behavior.

3.2 Contractible reliability

As a benchmark for the following analysis, consider now a (hypothetical) scenario where reliability and the rewards for reliability are not only observable, but verifiable as well. Thus, reliability is *contractible*. Because the manager proposes the contract, she can simply ask for reliability and offer to pay the worker an amount x_c each period, where the index c denotes the contractible case. These reliability costs add to the wage bill. Thus, assuming for the moment that M indeed wants to induce reliability, in order to identify the optimal competence level, the manager maximizes

$$\max_{\theta} U_M = \underline{y} + \theta \Delta y + v(\theta) - (w + x_c) \tag{5}$$
$$s.t.\ U_W = w + x_c \geq \bar{u}(\theta).$$

Naturally, we have $x_c = b(\theta)$ each period.[4] The constraint indicates that inside utility, given by wages plus rewards for reliability, must, on average, be no less than reservation utility. To minimize wage costs, M will again offer $w = \bar{u}(\theta) - b(\theta)$. Thus, we see that the manager's objective is to

$$\max_{\theta} -\theta\alpha + v(\theta). \tag{6}$$

Because $\alpha < 0$, maximal competence $\theta = 1$ is optimal, leading to a payoff of $v(1) - \alpha$.

Comparing the payoffs under contractible reliability ($v(1) - \alpha$) and spot interaction ($b(1) - \alpha$), we have the intuitive

Proposition 1 *When reliability is contractible and has positive social (joint) value, the manager prefers it to the equilibrium that arises out of statically optimal behavior (spot interaction).*

A contractible reliability (static multi-tasking, contractible effort) model with positive correlation between inside and outside productivity therefore holds no predictions for the optimal level of competence except those stemming from comparative technology advantages

[4] This is equivalent to paying the expected present value of side opportunities to W up front.

(which translate into wage savings). This situation changes drastically when we relax the assumption that reliability is contractible.

3.3 Noncontractible reliability

In many cases, neither allegiance nor rewards for reliability can be proven in a courtroom. Reliability is non-verifiable and, thus, non-contractible. In this case, Table 1 shows that reliability cannot be sustained in the one-shot game. Because $\theta \geq 0, E_\theta y \geq 0, w \geq 0, x \geq 0$, and $b(\theta) \geq 0$, the unique Nash equilibrium of the stage game is spot interaction. Indeed, note that the game is similar to an asymmetric Prisoner's Dilemma. Repeated interaction is required to open up the possibility for M and W to obtain the reliability outcome. The crucial difference of this analysis from a standard repeated Prisoner's Dilemma is that M can choose which amount x to offer and with whom to play the game.

As regards timing, we need to ask whether M can take $v(\theta)$, but not pay x, and whether W can take x, but still do $b(\theta)$. That is, can the parties deviate from the reliability with rewards contract and still obtain the other side's cooperative contribution? In many practical situations, much like in the standard Prisoner's Dilemma used to analyze many types of social interactions, the answer to this question is yes. Even when M and W do not literally move simultaneously, this is the correct assumption to make when they do not learn the other party's move until later in the period, as seems plausible in many real-world circumstances.

Alternative assumptions are conceivable in some situations, and the analysis below also explores the implications of such alternative assumptions. For example, in some cases it is possible for M to pay a conditional bonus for W's acting reliably *in the same period.* The latter assumption is typically used when the focus of the analysis is on the incentive scheme that M can use to induce non-contractible effort ("pay for performance"). Essentially, this assumption amounts to making reliability contractible for one party and thus assuming that W has no commitment problem. (In this case then, M captures all the surplus. In the presently analyzed main case, by contrast, M needs to leave some surplus to W.) Results below will show that as long as M has a commitment problem (that is, as long as M cannot definitively promise to reward reliability after it has been delivered), a tradeoff between reliability and competence will still arise even if W can commit to reliability.

3.3.1 Non-reneging constraints and the feasibility of reliability

This section presents the conditions for reliability with rewards to be an equilibrium supported by trigger strategies. As is typical in the literature on repeated games, the analysis here focuses on the equilibrium that arises only under repeated play, but does not answer the question when this is a likely outcome. Specifically, consider self-enforcing (relational) *reliability with rewards* contracts. We concentrate on stationary contracts of the following form: W promises to be reliable to the manager. The manager promises to pay x in each period. Any one-time deviation by any player results in both players exerting the statically optimal behavior in all future periods.[5]

For the worker, honoring the promise of reliability means foregoing $b(\theta)$, but obtaining x in addition to the basic wage w in each period, thus obtaining

$$w + x + \frac{\delta}{1-\delta}(w + x) \tag{7}$$

in total welfare. By deviating today, W obtains extra utility $b(\theta)$ today but suffers a reduction to the reservation utility going forward (which comes from spot interaction forever, or, equivalently for him, a break-up of the relationship). Thus, by deviating today, W obtains

$$w + b(\theta) + x + \frac{\delta}{1-\delta}(\bar{u}(\theta)) \tag{8}$$

in total welfare. Recall that $\bar{u}(\theta)$ is equal to $w + b(\theta)$.[6]

[5] It is assumed that M does not hire a new W (with a possibly different competence level). Abreu (1988) showed that if cooperation is attainable in a repeated game, it is without loss of generality to concentrate on the worst punishment path. As shown above, spot interaction yields the same utility for the worker as if no relationship is formed, making it the worst punishment path. The approach here has become the standard approach for the analysis of self-enforcing contracts and has been used both in methodological research (Bull (1987), MacLeod and Malcomson (1989), Levin (2003)) and in applications such as Baker, Gibbons, and Murphy (2002). There would be state-varying payments if the outside wage were also state-varying (Thomas and Worrall 1988). For recent work on non-stationarity contracts (due to limited liability or other frictions) see, for example, Thomas and Worrall (2010) and Yang (2009).

[6] M does not take away W's side opportunities on the reneging path. To see why this makes sense, recall that M saves spot wages through this – she has to convey utility $\bar{u}(\theta)$ to W either way. If M could take

Therefore, combining (7) and (8), the worker's non-reneging constraint (NR-W) is

$$w + x + \frac{\delta}{1-\delta}(w+x) \geq w + b(\theta) + x + \frac{\delta}{1-\delta}(w + b(\theta)). \tag{9}$$

We next see that the spot wage (the fixed component of the remuneration) drops out.[7] Simplifying yields

$$\delta x \geq b(\theta). \tag{10}$$

In other words, the required reliability rewards are the expected value of side opportunities scaled up by the inverse of the discount factor. Let

$$\tilde{x}(\theta) = b(\theta)/\delta \tag{11}$$

denote the *minimal rewards that induce reliability from type θ.*

In addition, the relational contract that M offers must also satisfy W's participation constraint

$$w + x \geq \bar{u}(\theta). \tag{12}$$

But any worker who is in spot interaction with M earns $w = \bar{u}(\theta) - b(\theta)$. M would not pay a higher base wage to a reliable worker but would use the loyalty rewards x as a bonus. Plugging in, (12) becomes $\bar{u}(\theta) - b(\theta) + x \geq \bar{u}(\theta)$, which simplifies to the participation constraint

$$x \geq b(\theta). \tag{13}$$

Because the self-enforcing contract needs to create surplus for the reliable worker, (13) is implied by (10), W's non-reneging constraint. That is:

Lemma 1 *Whenever the worker's non-reneging constraint is fulfilled, his participation constraint is also satisfied.*

away W's side opportunities and pay him less, spot interaction would cease to be the worst punishment path, inconsistent with the standard modeling approach.

[7] That the wage is a transfer and drops out from the non-reneging constraint is a fact common to models of relational contracts with additively separable fixed and bonus components of remuneration. The same occurs, for example, in MacLeod and Malcomson (1998).

Thus, in principle, every worker can be induced to be reliable – as long as the rewards x are high enough. But that is precisely the constraint the manager faces. The higher the promised x, the higher is the temptation to renege on the contract. For some realized output today, M's non-reneging constraint requires that obtaining reliability benefits $v(\theta)$ net of reliability rewards x in each period is preferred to receiving reliability today without paying for it, but never receiving nor paying for reliability in the future. Therefore, the manager's non-reneging constraint (NR-M) is

$$y + v(\theta) - x - w + \frac{\delta}{1-\delta}\left[E_\theta y + v(\theta) - x - w\right] \geq y + v(\theta) - w + \frac{\delta}{1-\delta}\left[E_\theta y - w\right] \quad (14)$$

which collapses to

$$x \leq \delta v(\theta). \quad (15)$$

In other words, the maximum credible reliability rewards are the value of reliability to the manager scaled down by the discount factor.

The manager will set x as small as possible, but she needs to take into consideration that NR-W is still fulfilled. When the parameter values are such that both constraints can hold, there are (infinitely) many divisions of the surplus that work (see Theorem 1 in Levin (2003)). Even without taking a stance on the division of this surplus, however, we can see that whether reliability dictates less than full competence – and whether reliability is feasible at all – depends on the shape of the two critical functions, $b(\theta)$ and $v(\theta)$.

Formally, Proposition 2 summarizes the results by combining the two non-reneging constraints. The Corrolary follows from the fact that contractible reliability is desirable when $v(\theta) > b(\theta)$.

Proposition 2 ***(Feasibility of reliability with rewards)*** *Reliability is feasible at some given competence level θ^R if*

$$\frac{b(\theta^R)}{v(\theta^R)} \leq \delta^2. \quad (16)$$

Corollary 1 *Non-contractible reliability may not be feasible, even when reliability would be desired if it were contractible.*

This is a simple and intuitive set of results. Reliability can only be obtained where M can credibly threaten W with severe punishments if he is not reliable. One key determinant of M's punishment ability is W's competence, and how competence affects side opportunities and inside value generated through reliability. Note that M's relative productivity α has no implications for the feasibility of reliability (although it will have implications for its desirability).

To develop intuition further, Figure 1 depicts the situation graphically, using examples of different assumptions about $b(\theta)$ and $v(\theta)$ in the different panels. (The Figure plots $v(\theta)$ concavely and $b(\theta)$ linearly, although this is not by any means a required condition for the model to apply. Also, the Figure is not a taxonomy of all possible cases.)

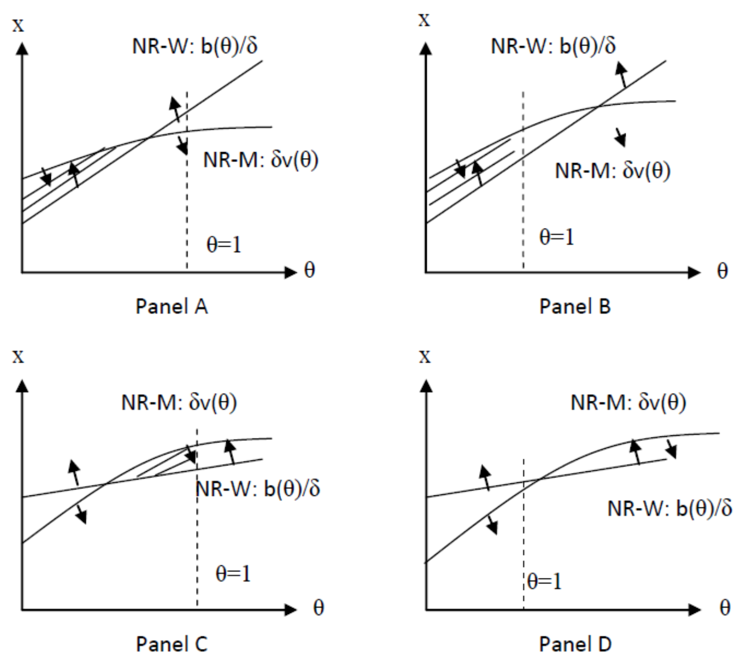

Figure 1: The feasibility of reliability. Competence θ can be 1 at most, but for comparison across the panels, it is instructive to see the shape of the non-reneging constraints extend over a broader range. NR-W is the worker's non-reneging constraint: Only reliability rewards above this line are sufficient for the worker to behave reliably. NR-M is the manager's non-reneging constraint: Only reliability rewards below this line can be credibly promised. The shaded area shows the feasible set. The Figure shows four examples, not a taxonomy. (The text discusses all cases.) A tradeoff between reliability and competence arises in the case of Panel A. Reliability is feasible for all competence levels in Panels B. Reliability is feasible only above a certain competence level in Panel C, and for no competence level in Panel D.

Panel A in Figure 1 depicts one arguably plausible benchmark case. Here, incompetent workers create more value through reliability than they have outside options ($v(0)\delta > b(0)/\delta$). And although the value of reliability is greater the more competent the worker is ($v' > 0$), the value of outside options increases sufficiently fast so that the two non-reneging constraints cross at some $\theta < 1$. Reliability is only feasible in the shaded region, i.e., only for those agents with sufficiently low levels of competence. If the value of reliability increases faster and/or side opportunities increase more slowly and/or the surplus generated by reliability of those with a zero competence level is high enough, full competence may be compatible with reliability (see Panel B). If reliability is infeasible for incompetent workers because they have relatively good outside options, reliability may be feasible for only sufficiently high levels of competence (Panel C) or no competence levels (Panel D).

Thus, even at high competence levels, reliability may be feasible if the value added created through increases in competence is sufficiently large. On the other hand, a tradeoff between reliability and competence is likely to arise if incompetent agents produce sufficient value of reliability and if competence matters more outside than inside, taking discounting into account ($b'/\delta^2 > v'$).

This also implies that there can be a tradeoff between reliability and competence that M needs to optimize on. If the maximum θ^R for which condition (16) holds is less than unity, M must choose whether to hire a fully competent W who will not behave reliably, or W who will act reliably, but who is less competent. The next subsection analyzes this tradeoff and the resulting comparative statics for the optimal competence level.

3.3.2 The optimal competence level

When implementing reliability, M chooses W's optimal competence by maximizing his expected utility, $E_\theta y + v(\theta) - x - w$, subject to the constraint that reliability must be feasible. Recall that $w = E_\theta y + \alpha\theta - b(\theta)$. From equation (11), the loyalty rewards that M will pay are $x = b(\theta)/\delta$. Plugging in, M's optimization problem therefore is

$$\max_\theta f(\theta) = -\alpha\theta + b(\theta) + v(\theta) - \frac{b(\theta)}{\delta} \tag{17}$$

$$s.t.\ \frac{b(\theta)}{v(\theta)} \leq \delta^2,$$

with θ between 0 and 1. Allowing $b(\theta)$ and $v(\theta)$ to take on completely arbitrary shapes complicates further analytical statements tremendously; a little bit of structure, drawing on commonly used functional forms, helps clarify matters. In particular, suppose that the reliability value produced by competence follows a concave, power utility/production function with a shift term, $v(\theta) = (1+\theta)^\gamma + V$, where $V > 0$, $0 < \gamma < 1$. Also, suppose that $b(\theta) = \beta\theta + B$, where $B > 0$, $\beta > 0$. These functional forms are fairly common.

To understand M's choice, it is helpful to note the conditions when she would like to maximize worker competence. If $-\alpha > \frac{1-\delta}{\delta}\beta$, then $f'(\theta) = \left(-\alpha - \frac{1-\delta}{\delta}\beta\right) + \gamma(1+\theta)^{\gamma-1} > 0$ $\forall\theta > 0$. Intuitively, when M's productivity advantage (as measured by $-\alpha$) is about as big as the marginal product that W's competence has for the search for side opportunities (as measured by β), then the principal's objective function is monotonically increasing in W's competence (when δ is greater than one-half). As δ increases (as players get more patient), even when M has a smaller relative productivity advantage (a smaller $-\alpha$) or W's competence gives rise to many additional side opportunities (a larger β), M would like to hire the most competent worker possible. Therefore, this case is quite likely to occur and is, therefore, the main case of interest studied here. (See below for a treatment of other cases.)

Case 1: There is an upper bound on the level of competence compatible with reliability. In the concrete setting, this case arises when $V + 1 \geq \frac{B}{\delta^2}$, i.e., when incompetent workers add more reliability value than they have in terms of side opportunities, even adjusted for discounting. (Feasibility of reliability also implies social efficiency.) Moreover, the interesting case occurs when there exists an intersection point of $b(\theta)$ with $v(\theta)$ in $[0, 1]$. This is the case of panel A in Figure 1.

Here, we obtain the following result and comparative statics:

Proposition 3 *Suppose that there is an upper bound on the level of competence compatible with reliability and suppose that $f'(\theta) > 0$. The optimal θ^{R*} has to fulfill $\frac{\beta}{\delta^2}\theta^{R*} - (1+\theta^{R*})^\gamma = V - \frac{B}{\delta^2}$. For sufficiently small changes of δ, B, β, V, and γ such that all assumptions are satisfied, we obtain that: The optimal level of competence compatible with reliability is higher,* *the more patient the manager and the worker are, the worse side opportunities are, the less side opportunities vary with worker competence, the more important reliability is, and the faster the value of reliability increases with competence. That is,* $\frac{\partial \theta^{R*}}{\partial \delta} > 0$, $\frac{\partial \theta^{R*}}{\partial B} < 0$, $\frac{\partial \theta^{R*}}{\partial \beta} < 0$, $\frac{\partial \theta^{R*}}{\partial V} > 0$, $\frac{\partial \theta^{R*}}{\partial \gamma} > 0$.

Proof. See the Appendix. ■

It is intuitive that as patience (δ) increases, θ^{R*} increases. As is well-known from the Folk-Theorems, more patient individuals are less likely to renege on their promise or, conversely, the temptation $b(\theta)$ weighs comparatively less. We would, therefore, expect project-based organizations – which have a shorter time horizon – to have difficulties in acquiring competent and reliable partners for projects.[8]

Next, environments where greater competence affords significantly better side opportunities for W (higher β) or which hold better side opportunities for any level of competence (higher B) tend to bring about a tradeoff between competence and reliability. In a small alpine village, hiring the most competent person does not cause great concerns for lack of reliability, whereas in New York City the tradeoff between competence and reliability should be more pronounced.[9]

Finally, note that as the degree by which competence makes reliability more valuable rises (γ increases) or as reliability itself simply becomes more important (V increases), even non-contractible reliability can allow full competence. One might think that the tradeoff between competence and reliability implies that where reliability is more important (i.e., where $v(\theta)$ is larger), there will be less competence. But in fact, the opposite is true in a sense. The reason is that as the value of reliability increases, the manager can now more credibly promise reliability rewards. This result has important implications for the choice of optimal competence, to which the analysis turns next.

To make statements about the overall optimal level of competence, we can compare the welfare that M obtains in reliability, $f(\theta^{R*})$, with the welfare from optimally solving the spot

[8] However, there may be recurring projects. For example, although the film industry is very much characterized by projects, reliability plays a role because the players expect to meet again.

[9] A related result is that of Ramey and Watson (2001) who find that a fall in market friction, as reflected by a rise in the probability of locating a new trading partner in the matching market, makes less onerous the threat of severing an existing trading relationship, thereby tightening the effort constraint.

interaction problem, $-\alpha + b(1)$. The following proposition summarizes the results.

Proposition 4 *Suppose that there is an upper bound on the level of competence compatible with reliability. The manager prefers reliability (with less than full competence) to spot interaction (with full competence) if the optimal level of competence compatible with reliability is high enough. (The threshold level of competence compatible with reliability is derived in the Appendix.)*

Proof. See the Appendix. ■

Corollary 2 *There is a non-monotonic relationship between observed equilibrium competence and the factors that drive the optimal competence in reliability.*

To see this, note first that competence in reliability is always at most as high as in spot interaction. Second, some differences between the variables need to be considered that determine how the details work out. The marginal productivity of competence for the value of reliability (γ) and the value of reliability that incompetent workers generate (V) only affect the principal's welfare if she does, indeed, implement reliability. Consider a case where M and W interact in spot interaction. An increase in the value of reliability (either γ or V) increases the level of competence compatible with reliability. This may make reliability the preferred outcome, in which case optimal competence suddenly falls. Within reliability equilibrium behavior, however, further increases in the value of reliability enhance optimal compentence (see Proposition 3), overall yielding a non-monotonic relationship between the value of reliability and the level of agent competence the principal prefers.

These comparative statics distinguish this model from existing work. In the relational contracts model of Baker, Gibbons, and Murphy (2002), when reliable delivery by the upstream party of a good to the downstream party becomes more important, making the upstream party more "competent" by way of non-integration (that is, giving it a better outside option) becomes less attractive. In contrast to asset ownership, competence is by nature continuous, allowing for the richer comparative statics.[10]

[10] Of course, this is not to say that a relational contracts model is needed to obtain a non-monotonic

Conversely, starting in reliability, an overall higher level of side opportunities, B, reduces the feasible (and, within a range, optimal) competence level, makes reliability less preferred, and ultimately leads the principal to switch to independence with maximal competence, again yielding a non-monotonic relationship overall. For the extent to which side opportunities for the worker depend on worker competence, β, the result is similar but further complicated by the fact that both the threshold for reliability to be optimal (formally derived in the Appendix) and θ^{R*} itself depend on β. On the one hand, if β increases, spot interaction becomes more attractive. The optimal level of competence under reliability also decreases, making reliability less attractive. On the other hand, M saves more in wages for each given competence level. Therefore, M may, in fact, be more likely to implement reliability. Interestingly, therefore, the model implies that reliability inside companies does not necessarily have to decrease if W can use his competence in more ways in side opportunities. Similarly, while greater patience of the players (higher δ) allows a higher level of competence to be compatible with reliability, the present value of reliability reward costs also increases, making it possible for spot interaction to become more attractive as δ increases.

Finally, greater productivity of M ($-\alpha$) leads M to more likely implement spot interaction and to hire the most competent worker.

For an additional perspective on the results, denote the *net value of reliability* for a given type by $v(\theta) - \tilde{x}(\theta)$. This definition does not take into account M's non-reneging constraint. In other words, even if there is a positive net value of reliability, M may not necessarily be able to implement it. However, when reliability is feasible, we know that it has positive net value, i.e., $v\left(\theta^{R*}\right) - \tilde{x}\left(\theta^{R*}\right) = v(\theta) - b\left(\theta^{R*}\right)/\delta > 0$. Therefore, we have the intuitive

Corollary 3 *Even when reliability is feasible and has positive net value, the manager may prefer spot interaction with a more competent worker.*

relationship between optimal agent competence and the value of reliability to the principal. The focus here is on what the implications of the correlation between inside and outside opportunities are, given that the relational contracts model is the framework of interest. Having said that, the model does yield different imlications than some static models that study a tradeoff between loyalty and competence of the agent. For example, in the static model of Egorov and Sonin (2011), the more important the loyalty of the vizier, the more the dictator is inclined to choose an incompetent vizier.

Case 2: There is a lower bound on the level of competence compatible with reliability. This is the case of panel C in Figure 1. This case arises when $B \geq V+1$ (that is, incompetent W's have side opportunities that have greater value than what reliability is worth to M, i.e., reliability is socially inefficient) and when there exists a unique intersection point of the two lines in $[0,1]$ which is in $(0,1)$. We immediately have:

Proposition 5 *Suppose that there is a lower bound on the level of competence compatible with reliability. The manager will always choose a maximally competent worker under reliability,* $\theta^{R*}=1$. *That is, for sufficiently small changes of* δ, B, β, V, *and* γ *such that all assumptions are satisfied,* $\frac{\partial\theta^{R*}}{\partial\delta}=0$, $\frac{\partial\theta^{R*}}{\partial B}=0$, $\frac{\partial\theta^{R*}}{\partial\beta}=0$, $\frac{\partial\theta^{R*}}{\partial V}=0$, $\frac{\partial\theta^{R*}}{\partial\gamma}=0$.

Even though it is optimal for M to have a maximally competent worker, equilibrium behavior of the worker will depend on the parameters. Specifically, M prefers to implement reliability rather than spot interaction if

$$f(1)=-\alpha+2^{\gamma}+V-\frac{1-\delta}{\delta}(\beta+B)>-\alpha+b(1)=-\alpha+\beta+B. \tag{18}$$

Not surprisingly, the higher V and γ, the more likely it is that $f(1)>-\alpha+b(1)$, making reliability the preferred strategy; and the higher B and β, the more likely it is that spot interaction is the preferred outcome.

Other cases. The other cases in Figure 1 are straightforward, but do not yield interesting comparative statics. First, when $b(\theta)$ and $v(\theta)$ do not cross in $[0,1]$ as in panel B, maximal competence is both feasible and optimal (and reliability is always desired). Second, if reliability is infeasible for any competence level, as in panel D, M will always implement spot interaction with maximal competence.

There are also cases not covered in Figure 1. For example, when $b(\theta)$ and $v(\theta)$ cross twice in $[0,1]$, competence may be infeasible up to a certain competence level and above another competence level, but feasible in the intermediate range. When M's objective is increasing in W's competence, this case is similar to case 1 discussed above. If the principal's welfare is not monotonically increasing in competence, then three cases need to be distinguished, see Proposition A.1 in the Appendix.

3.3.3 A useful special case with an explicit solution

Consider the case when the value created through reliability does not vary with competence, i.e., where $v(\theta) = v$. In a majority of cases, one would have to view this as an unrealistic assumption. However, from this case further interesting results can be obtained that make it worthwhile paying the price of lack of realism.[11]

To simplify matters further, assume that $b(\theta) = \theta b$. (Equivalently, the outside options are such that in each period, a side opportunity of value $b > 0$ arises for W with θ, so that the expected value of side opportunities in each period is θb.) To ensure that M does not want to hire a W of negative competence, assume that $\frac{\delta^2}{1-\delta} \geq \frac{b}{v}$. These assumptions together describe the *case of a constant value of reliability.*

Proposition 6 *Consider the case of a constant value of reliability. Reliability is feasible if and only if*

$$\theta \leq 1 - \frac{b - \delta^2 v}{\delta b} = \theta^{R*}. \tag{19}$$

Even when reliability is feasible and has positive net value, the manager may prefer spot interaction with a more competent worker. The more productive the manager is (i.e., $(b - \alpha)$ is large) or the tighter the bound on the feasibility of reliability in terms of the allowed competence levels is (i.e., $\left(1 - \theta^{R*}\right)$ is large), the more valuable reliability must be to the manager in order to be preferred over spot interaction, where full competence is feasible. The empirical predictions parallel those of the general model.

Proof. See the Appendix. ■

In this case, therefore, a tradeoff between reliability and competence is inevitable. Recall that this result is derived assuming that both M and W have a commitment problem in the sense that the move of M (W) is not seen by W (M) before M (W) makes a decision. However:

[11] Although the assumption that v is constant is strong, this case may approximately cover two scenarios. First, an act of reliability may literally involve attending a vote and/or casting the right vote in a democratic decision process, e.g., a committee or a board in an organization, rather than taking care of W's own business. (This neglects that the value of W's vote will depend on the votes of others.) Second, W's decision may be approving a project, the authority over which M has delegated to W (and where the principal does not immediately break up the relationship in case of violation of the principal's wishes).

Proposition 7 *Consider the case of a constant value of reliability. A necessary and sufficient condition for the tradeoff between reliability and competence to arise is that the manager cannot commit to reward the worker for reliability.*

Proof. See the Appendix. ■

A lack of commitment is a necessary condition for the trade-off to exist even in the general case. (If M could commit, the contract would be explicit and reliability could always be achieved.) What the Proposition says is that even when we make stronger assumptions that in principle make the tradeoff likely (in particular, when v does not increase in θ), there will not be a tradeoff if M is not tempted to renege on her promise, but if he is tempted, then this is sufficient for the trade-off to arise even if W has no commitment problem. This is important in that some managers or organizations may be able to develop a reputation for rewarding reliability, and may thus not be constrained by a non-reneging temptation on their part.

4 Applications and evidence

The basic model insight relevant for many real-world situations is that a relational contracts model with a positive correlation of inside and outside productivities tends to yield a trade-off between reliability and competence.[12] Conditional on a relational contracts model being the appropriate model framework for a given setting,[13] allowing for the correlation is attractive

[12] In principle, one needs to determine which of the cases discussed in the model analysis applies on a case-by-base basis. By and large, it seems plausible that the tradeoff can arise within organizations. For example, a more competent computer expert at an internet company may be more apt at applying his superior knowledge to the particular tasks at hand, and the value generated for his manager by reliable acts may also be greater. However, his side opportunities probably increase much more strongly with his competence than the value of reliability inside the relationship. Not only can he engage in other similar programming tasks, but knowing how to program in a computer language is a universally applicable tool. Thus, even this relatively specific skill makes side opportunities grow much faster with competence than inside contributions do. The expert can also do general consulting about the software business. Note that this is a statement about the relative marginal productivities of competence in the non-contractible parts of production, $v(\theta)$ and $b(\theta)$. We may well have $\alpha < 0$ at the same time, which holds that in terms of the contractible parts, the match between M and W makes better use of W's competence than W's alternative employment.

[13] As discussed in the introduction, other model frameworks can also generate the prediction that managers may not wish to hire the most competent employee.

because the present framework implies that the principal will not choose the person with the least valuable side opportunities (as she does in the absence of the correlation), but will balance the concern for relations against the concern for productivity.

Generally speaking, significant evidence exists that is consistent with the model's overall message, but, as we will see, few direct tests of unique predictions of the model are yet available. The available empirical evidence regarding the model comes in three parts.

First, more competent individuals in fact tend to behave less reliably. For example, Williams (1996) discusses the problems managers face with very intelligent software developers who are less easily induced to follow their managers' plans. Moreover, studies in organizational behavior have found that those with many outside options (an important determinant of which is competence) are less likely to act loyally and to contribute organizational citizenship behavior (Bergeron 2007, Lawler and Yoon 1996, Podsakoff, MacKenzie, Paine, and Bachrach 2000).

Second, principals recognize the trade-off in practice. The management literature routinely points to the practice of managers hiring less than fully competent workers to ensure reliability (Sample 2003). Samuel Goldwyn famously stated: "*I'll take 50% efficiency, if I can get 100% loyalty.*" Consistent with this view, Larry Ellison's hiring practices for his senior executives at Oracle and Richard Auhull's choices of allies for Circon (Hall, Rose, and Subramanian 2001) have been interpreted as favoring reliability over competence. Also, chairmen of boards, when they "move up" from the CEO's chair, do not always choose the most competent CEO as their successor but seek to appoint a loyal individual.[14] Managers also frequently do not choose the most competent advisors because they fear they will be less reliable. For example, Joni (2004) describes the case of a regional vice president who was worried that his subordinates knew the territory and the players too well and "feared they would exploit that knowledge for their own purposes" (p. 87). Examples abound also in political economy; reliability of executives in public agencies is valued highly. Even in the

[14] Conversely, a simple version of the model was applied in Wagner (2011) to investigate why CEOs may not desire the most competent board members (though this simpler model does not allow for several features shown to be important in the more general framework here, such as the non-monotonicity of optimal agent competence).

US and other developed countries, many "agency executives are selected in order to serve the political needs of the president, and these may or may not involve policy considerations" (Wilson 1989, p. 198). The presidential appointment process is frequently regarded as being strongly dominated by the tradeoff between competence and reliability (Edwards 2001), sometimes leading to "amateur government" (Cohen 1998). Similarly, Heclo (1977) observed that "many of [the agency's executive's] selectors [are] more interested in the process of getting their way than in the executive's eventual output" (p. 99).

Third, while a model of relational contracts that takes into account the impact of productivity provides one way to organize management and political folklore into a coherent picture, the cited evidence does not in fact constitute a test of the more specific features of the present model. Indeed, limited evidence of this sort is currently available. The model predicts that (in the range where reliability is optimal) firms with a higher discount factor will hire more competent workers. To the extent that larger (more mature and older) firms have a higher discount factor, the model is thus one possible explanation for the stylized fact of the positive correlation between firm size and wage level. (Of course, there are also other explanations, such as a larger impact of worker skill on output in larg firms compared to small firms.) Wagner (2010) empirically considers the cross-section of competence levels in public agencies. The analysis reveals that recruitment into public agencies is more meritocratic and more focused on competence in countries where (1) agency officials have poorer outside options (for example, because they have few private sector contacts), (2) careers in agencies are expected to be longer-lasting, and (3) agencies are more powerful (and, thus, their reliability arguably more important). To the extent that the observed agency-government relations capture reliability with rewards equilibria, this evidence is consistent with Propositions 3 and 6. By contrast, finding (3) in particular is not consistent with alternative theories, such as Egorov and Sonin (2011). While the Wagner (2010) study appears to be the only empirical analysis available that tests some of the relational contracts framework's implications, it also does not test the prediction regarding the non-monotonic relationship between importance of reliability and agent competence as it assumes that the agencies and governments are playing according to an implicit reliability contract. A full test of the present theory would require data covering both reliability equilibria and spot interactions. Future research can hopefully make progress on such a test.

5 Extension: Signaling with reliability

We have so far assumed perfect and symmetric information. In reality, the manager or the market may not know a worker's competence with the same precision as he does. This section makes a first step towards extending the model to include the possibility that agents signal their competence through independence or reliability, and it suggests that this model framework may help explain some stylized facts regarding labor markets.

To model this situation, for tractability, consider the simplified model with a constant value of reliability. Suppose that M has one worker who can be of two different types, High and Low, with levels of competence $0<\theta_L \leq \theta_H < 1$. Begin by assuming that M knows the worker's type, but the market does not know which type the worker is. The manager chooses the wages she offers to the high and low types, respectively, w_H, w_L and reliability rewards x.[15]

At the beginning of each period, conditional on the history of play, the market updates its belief (starting from some prior) about the worker's competence. Consider the simplest case, where in the first period an opportunity of value b arises for sure, and the market knows this. Thus, the market learns whether W behaved independently or reliably. This implies that if the market believes that only Lows behave reliably, separation occurs either immediately or never. It is assumed that the market does not observe wages or output levels. This is a strong assumption; it is an extreme version of the perhaps palpable notion that the market is more likely to observe external actions than company-internal variables.

In a perfect Bayesian equilibrium, given the market's beliefs and the manager's wages and rewards schedule, both worker types find it optimal to act according to the market's belief. Given the market's beliefs and the workers' behavior, the manager finds it optimal to offer reliability rewards and wages consistent with beliefs and behavior.

[15] In principle, the manager can offer two rewards levels, x_H and x_L, but in the equilibrium we consider, it will turn out to be optimal to induce only θ_L to reliability.

Analysis If M wants less competent workers to behave reliabily, the potential problem is that if the market's rewards for independence are too tempting, low-competence workers will try to mimic high-competence workers. Begin by assuming that even when M knows the worker is Low, in order to keep him inside the firm, she would need to offer a wage corresponding to High's wage forever if W acts independently. Despite this strong assumption, separation is nonetheless potentially possible in equilibrium. The tradeoff between competence and reliability provides for an endogenous sorting condition: reliability is less attractive for High than for Low. In other words, the implicit costs of reliability instead of no reliability are higher for High than for Low.

Employing these insights, and recognizing that reliability payments also need to cover the wage differential an independently behaving worker can obtain, we can state the following result for the *basic signaling game* $\Gamma(\delta, b, v, Y, \{\theta_H, \theta_L\})$ with the following features: (1) An opportunity of value b appears for certain at the beginning of the game. (2) The market values reliability less than M and $\alpha = 0$. (3) Assume $\bar{y} > b$ and $\theta_H - \theta_L > \frac{1-\delta}{\delta}$. (4) Low's reliability has positive net value under perfect information (i.e., $v - \tilde{x}(\theta_L) > 0$) and High's reliability has negative net value under perfect information (i.e., $v - \tilde{x}(\theta_H) < 0$).

Proposition 8 *Consider the basic signaling game. There exists a separating perfect Bayesian equilibrium that is characterized by the following features:*

1) The market believes that θ_H acts independently and that θ_L acts reliably. The two types behave accordingly.

2) The reliability rewards x and wages w_H, w_L satisfy the following conditions:

a) High is at least as well off as Low. That is: $w_H + \theta_H b \geq w_L + x$.

b) Low does not want to be independent. That is: $w_L + x > w_H + \theta_L b$.

c) High exactly receives his reservation utility. That is: $w_H + \theta_H b = \bar{u}(\theta_H)$.

d) Low earns a surplus over his reservation utility. That is: $w_L + x > \bar{u}(\theta_L)$.

(3) The manager never does better than if the market has symmetric information.

Proof. See the Appendix. ■

In other words, whenever M wants to separate types under perfect information, she can do so under asymmetric information (as long as the types are sufficiently different and sufficiently patient), but at additional costs. Note that the characterization of wages and reliability rewards in part (2) allows for a range of equilibria. This multiplicity of equilibria is not due to a problem of specifying beliefs on an off-equilibrium path, but due to the fact that wages and reliability payments, $w_L + x$, are essentially perfect substitutes in inducing Low to reliability.[16]

Discussion Four comments on these results are in order. First, we had assumed that Low is able to secure w_H forever by being independent. This similarity between Low and High is arguably too extreme. To the extent that Low can only achieve a lower wage, separation between types is easier.

Second, recall that Proposition 8 is valid for the case where other firms care less about reliability than M. If the opposite is true, Low has less incentive to mimic High. The reason is that signaling high competence now is in fact a bad signal - no one wants to appear *overqualified*.

Third, what if M does not know W's competence either? Even in this case, we can show that the same logic applies, i.e., that M may use the sorting possibility presented to her in the form of the reliability-competence tradeoff to her advantage.

Proposition 8' *Under the conditions of Proposition 8, a separating equilibrium with the properties of the equilibrium in Proposition 8 exists also in the case where the manager is asymmetrically informed about the worker's competence.*

Proof. See the Appendix. ■

Fourth, these results can be compared with Fryer (2007). In his paper, highly talented agents in a community do not invest in group-specific capital, while less talented individuals do invest in sufficient amounts in order to signal that they would like to engage in cooperative behavior with the target group. In his analysis, thus, a signal is conveyed through an

[16] Each type has exactly one behavior to exhibit in equilibrium and there are only two possible behaviors. Therefore, no specification of off-equilibrium beliefs is necessary. Thus, this equilibrium is also sequential (Kreps and Wilson 1982). As before, the reason for treating wages and reliability rewards separately is that M can only renege on reliability rewards.

investment before a cooperation game is played. In the present paper, the reliability-competence tradeoff gives rise to an endogenous screening condition, i.e., those who show independent behavior today while playing the game are inferred to be competent and thus predicted to behave independently in the future.

Evidence Because the case of a constant v is a limiting case of the more general instance $v = v(\theta)$, it appears likely that the results derived here also hold in a broader range of circumstances. To the extent that they do, they may help to inform a new interpretation of otherwise puzzling pieces of evidence from labor markets.

One implication of the model is that M cannot cut wages and reliability rewards for reliable workers too much. By "ratcheting" and offering too low a wage for the reliable worker, the manager risks that the worker is tempted too much by the ability to get the higher wage that must be offered to independently behaving workers. Thus, when we interpret wages and reliability rewards together and assume that when the manager has a choice she will pay out most of the sum in wages, not in informal reliability rewards, the model suggests one possible avenue towards a resolution to the puzzle that managers are generally reluctant to lower wages in the face of a decline in a worker's reservation wage.[17]

Second, the model can speak to the wage effects of mobility. The standard approach to analyze these effects is in terms of adverse selection. In particular, Greenwald (1986) shows that in this case workers who change jobs are marked by being part of an inferior group of competence (see also Spence (1973), McCormick (1990), and Gibbons and Katz (1991)). Light and McGarry (1998) and Munasinghe and Sigman (2004) indeed find that previous mobility is negatively related to wages and positively related to future turnover. But the explanatory variable they employ is total separations. Instead, the present model predicts that voluntary and involuntary separations have different effects. Voluntary separation tends to signal higher competence. Consistent with this prediction, Keith (1993) shows that once

[17] The literature frequently explains the fact that there is little ratcheting by "moral" reasons. Bewley (1999) points to motivators "having to do with generosity" (p. 431). MacLeod (2001) suggests that it can be explained by the threat of less effort by a worker whose wage is cut. The model here suggests a specific meaning for this claim, effort here being the honoring of a relational contract. Another explanation, also derived in a relational contracts model, stems from models where there are privately-known shocks to the opportunity costs of making promised payments (Englmaier and Segal 2012, Li and Matouschek 2012).

these two forms of mobility are disaggregated, a history of more frequent voluntary quits increases (log) wages, while a history of involuntary mobility decreases wages.[18]

6 Conclusion

This paper contributes to the theory of relational contracts. The methodological innovation is to incorporate a correlation between inside and outside productivities into the analysis. Several assumptions were necessary to obtain a set of tractable results; future work may be able to relax these assumptions (e.g., the exogenous assumption of how competence affects inside and outside productivity; or some of the functional form assumptions). To the extent that a relaxation of these assumptions retains the key intuition of the model framework, the substantive contribution of the analysis is to reveal that in many situations reliability, i.e., a relational contract that requires the agent to choose costly actions that benefit the principal in return for non-contractible rewards, may not be feasible for highly competent workers. The stronger and surprising result is that even when reliability is feasible and has positive social value, it may be dominated by spot interaction with more competent individuals. Contrary to what initial intuition might suggest, the principal's optimal choice of agent competence exhibits non-monotonicities with respect to various parameters of interest. For example, starting from a case where the agent behaves independently, as the value of reliability to the principal increases, optimal agent competence will first, at some point, fall (to achieve the now attractive reliability equilibrium) and then rise again (because the principal can more credibly promise rewards when reliability is more important to her). Moreover, the paper shows that reliable behavior can be a signal of low competence. However, when the value generated by reliability increases sufficiently strongly with competence, the tradeoff need not arise, nor will it when the manager can otherwise credibly promise rewards. Importantly, the theoretical and empirical implications of this paper arise only from the introduction of

[18] See Mincer (1986), Keith and McWilliams (1995), and Keith and McWilliams (1999) for further evidence on the effectiveness of employed job search. Models of turnover do not easily yield predictions about wage behavior (see Burdett (1978), Weiss (1984), where wage effects arise from mobility costs, Jovanovic (1979a), and others). Human capital theory predicts a negative relationship between job mobility and investments in job specific skills (Becker 1975, Jovanovic 1979b).

a positive correlation between inside and outside productivities into a framework of self-enforcing contracts; they are neither a feature of a static multi-tasking world (where a correlation between inside and outside productivities was seen to hold no predictions for the preferred competence levels) nor of the standard relational contract framework (where the prediction of existing theories simply is that contracts are most easily sustained with individuals who have low reservation utilities). This observation suggests that models of this sort could more broadly contribute to the analysis of settings where both production and relations matter for organizational success.[19]

[19] Besides addressing issues in labor markets and inter-firm relationships, the theory could be employed in the analysis of marriage (Becker, Landes, and Michael 1977) and of activities that demand a high degree of loyalty to the organization, such as churches or sects (Iannaccone 1992, Berman 2003).

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56(2), 383–396.

Aghion, P., and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29.

Austen-Smith, D., and R. G. Fryer (2005): "An Economic Analysis of 'Acting White'," *Quarterly Journal of Economics*, 120(2), 551–583.

Axelrod, R. (1984): *The Evolution of Cooperation*. Basic Books, New York.

Baker, G., R. Gibbons, and K. Murphy (2002): "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 117(1), 39–83.

Baker, G., R. Gibbons, and K. J. Murphy (1994): "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 109(4), 1125–1156.

Becker, G. S. (1975): *Human Capital*. NBER, New York.

Becker, G. S., E. M. Landes, and R. T. Michael (1977): "An Economic Analysis of Marital Instability," *Journal of Political Economy*, 85(6), 1141–188.

Bergeron, D. M. (2007): "The potential paradox of organizational citizenship behavior: Good citizens at what cost?," *Academy of Management Review*, 32(4), 1078–1095.

Berman, E. (2003): "Hamas, Taliban and the Jewish Underground: An Economist's View of Radical Religious Militias," *Mimeo*.

Bewley, T. F. (1999): *Why Wages Don't Fall During a Recession*. Harvard University Press, Cambridge, MA.

Bull, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102(1), 147–160.

Burdett, K. (1978): "A Theory of Employee Job Search and Quit Rates," *American Economic Review*, 68(March), 212–220.

Chassang, S. (2010): "Building routines: Leaning, cooperation, and the dynamics of incomplete relational contracts," *American Economic Review*, 100(1), 448–465.

Cohen, D. M. (1998): "Amateur government," *Journal of Public Administration Research and Theory*, 8(4), 450–497.

Doornik, K. (2006): "Relational contracting in partnerships," *Journal of Economics and Management Strategy*, 15, 517–548.

Edwards, G. C. (2001): "Why not the best? The loyalty-competence trade-off in Presidential appointments," *The Brookings Review*, 19(2), 12–16.

Egorov, G., and K. Sonin (2011): "Dictators and their viziers: Endogenizing the loyalty-competence trade-off," *Journal of the European Economic Association*, 9(5), 903–930.

Englmaier, F., and C. Segal (2012): "Unions, Communication, and Cooperation in Organizations," *Working Paper*.

Fehr, E., and A. Falk (2002): "Psychological foundations of incentives," *European Economic Review*, 46(4-5), 687–724.

Fong, Y.-f., and J. Li (2012): "Relational contracts, efficiency wages, and employment dynamics," *Mimeo*.

Friebel, G., and M. Raith (2004): "Abuse of authority and hierarchical communication," *Rand Journal of Economics*, 35(2), 224–244.

Fryer, R. G. (2007): "A Model of Social Interactions and Endogenous Poverty Traps," *Rationality and Society*, 19(3), 335–366.

Fuchs, W. (2007): "Contracting with repeated moral hazard and private evaluations," *American Economic Review*, 97(4), 1432–1448.

Gibbons, R., and L. F. Katz (1991): "Layoffs and Lemons," *Journal of Labor Economics*, 9(4), 351–380.

Glazer, A. (2002): "Allies as rivals: internal and external rent seeking," *Journal of Economic Behavior and Organization*, 48, 155–162.

Greenwald, B. C. (1986): "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53(3), 325–347.

Grossman, S., and O. Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.

Halac, M. (2012): "Relational contracts and the value of relationships," *American Economic Review*, 102(2), 750–779.

Hall, B. J., C. Rose, and G. Subramanian (2001): "Circon (HBS Case 9-801-403)," .

Hart, O. (1995): *Firms, Contracts, and Financial Structure*. Clarendon Press, Oxford.

Heclo, H. (1977): *A Government of Strangers: Executive Politics in Washington*. Brookings Institution, Washington, D.C.

Hirschman, A. O. (1971): *Exit, Voice and Loyalty*. Harvard University Press, Cambridge.

Iannaccone, L. R. (1992): "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives," *Journal of Political Economy*, 100(2), 271–291.

Joni, S. A. (2004): "The Geography of Trust," *Harvard Business Review*, (March), 82–88.

Jovanovic, B. (1979a): "Firm-Specific Capital and Turnover," *Journal of Political Economy*, 87(6), 1246–1260.

——— (1979b): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87(5), 972–990.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93(1), 1–47.

Keith, K. (1993): "Reputation, Voluntary Mobility, and Wages," *Review of Economics and Statistics*, 75(3), 559–563.

Keith, K., and A. McWilliams (1995): "The wage effects of cumulative job mobility," *Industrial and Labor Relations Review*, 49(1), 121–137.

——— (1999): "The returns to mobility and job search by gender," *Industrial and Labor Relations Review*, 52(3), 460–477.

Kreps, D. M., and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50(4), 863–894.

Lawler, E. J., and J. Yoon (1996): "Commitment in Exchange Relations: Test of a Theory of Relational Cohesion," *American Sociological Review*, 61(1), 89–108.

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

Li, J., and N. Matouschek (2012): "Managing Conflicts in Relational Contracts," *American Economic Review*, forthcoming.

Light, A., and K. McGarry (1998): "Job Change Patterns and the Wages of Young Men," *Review of Economics and Statistics*, 80(2), 276–286.

MacLeod, W. B. (2001): *Incentives in Organizations: An Overview of Some of the Evidence and Theory*Kiel.

——— (2003): "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 93(1), 216–240.

MacLeod, W. B., and J. M. Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57(2), 447–480.

——— (1998): "Motivation and Markets," *American Economic Review*, 88(3), 388–411.

McCormick, B. (1990): "A Theory of Signalling During Job Search, Employment Efficiency, and "Stigmatised" Jobs," *Review of Economic Studies*, 57(2), 299–313.

Mincer, J. (1986): "Wage Changes in Job Changes," *Research in Labor Economics*, 8A.

Munasinghe, L., and K. Sigman (2004): "A hobo syndrome? Mobility, wages, and job turnover," *Labour Economics*, 11, 191–218.

Podsakoff, P. M., S. B. MacKenzie, J. B. Paine, and D. G. Bachrach (2000): "Organizational Citizenship Behaviors: A Critical Review of the Theoretical and Empirical Literature and Suggestions for Future Research," *Journal of Management*, 26(3), 513–563.

Prendergast, C. (1993): "A Theory of "Yes Men"," *American Economic Review*, 83(4), 757–770.

Prendergast, C., and R. H. Topel (1996): "Favoritism in organizations," *Journal of Political Economy*, 104(5), 958–978.

Ramey, G., and J. Watson (2001): "Bilateral Trade and Opportunism in a Matching Market," *Contributions to Theoretical Economics*, 1(1).

Rayo, L. (2007): "Relational incentives and moral hazard in teams," *Review of Economic Studies*, 74, 937–963.

Sample, S. B. (2003): *The Contrarian's Guide to Leadership*. Jossey-Bass, San Francisco.

Spence, M. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87(3), 355–374.

Thomas, J. P., and T. Worrall (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies*, 55, 541–554.

——— (2010): "Dynamic relational contracts with credit constraints," *Working Paper*.

Wagner, A. F. (2010): "Loyalty and competence in public agencies," *Public Choice*, 146(1), 145–162.

——— (2011): "Board independence and competence," *Journal of Financial Intermediation*, 20(1), 71–93.

Weiss, A. (1984): "Determinants of Quit Behavior," *Journal of Labor Economics*, 91(June), 420–442.

Williams, M. J. (1996): "What You Can Learn from the People Who Manage the People Who Write Software," *Harvard Management Update No. U9608A*.

Yang, H. (2009): "Nonstationary relational contracts," *Working Paper*.

This comprehensive Appendix can be made available on a website upon publication.

Supplementary Appendix

Proposition 3 *Suppose that there is an upper bound on the level of competence compatible with reliability* and suppose that $f'(\theta) > 0$. *The optimal* θ^{R*} *has to fulfill* $\frac{\beta}{\delta^2}\theta^{R*} - (1+\theta^{R*})^\gamma = V - \frac{B}{\delta^2}$. *For sufficiently small changes of* δ, B, β, V, *and* γ *such that all assumptions are satisfied, we obtain that: The optimal level of competence compatible with reliability is higher, the more patient the manager and the worker are, the worse side opportunities are, the less side opportunities vary with worker competence, the more important reliability is, and the faster the value of reliability increases with competence. That is,* $\frac{\partial\theta^{R*}}{\partial\delta} > 0$, $\frac{\partial\theta^{R*}}{\partial B} < 0$, $\frac{\partial\theta^{R*}}{\partial\beta} < 0$, $\frac{\partial\theta^{R*}}{\partial V} > 0$, $\frac{\partial\theta^{R*}}{\partial\gamma} > 0$.

Proof of Proposition 3

It is helpful to define the following sets: $A := \{\theta : \frac{b(\theta)}{v(\theta)} \leq \delta^2\}$ and $B := [0,1] \cap A$. Additionally, to simplify notation, we define the linear function $l(\theta) := \frac{1}{\delta^2}(\beta\theta + B)$. Then,

$$B = \{\theta \in [0,1] : v(\theta) \geq l(\theta)\} \tag{20}$$

Note that B is connected or empty since $l(\theta)$ and $v(\theta)$ have zero, one or two intersection points between 0 and 1. In all three cases, the set of θs with $l(\theta) \leq v(\theta)$ is connected or empty in $[0,1]$ because v is strictly concave and hence, there is at most one region with $v(\theta) > l(\theta)$.

We also need the following definitions. Denote with θ^{R*} the optimal level of competence under reliability:

$$\theta^{R*} := \text{argmax}_{x\in B} f(x) \text{ if } B \neq \emptyset \text{ and argmax}_{x\in B} f(x) \text{ exists,} \tag{21}$$

$$q := \max B, \tag{22}$$

$$s := \min B. \tag{23}$$

When M's objective function is increasing in W's competence, the optimal level of competence is the maximal level of competence compatible with reliability, q.

In the present case, $B \neq \emptyset$. Per the assumption that M's welfare is increasing in W's competence, $\theta^{R*} = q$. As seen in panel A, the function $v(\theta)$ crosses $l(\theta)$ from above, and only competence levels lower than the intersection point are compatible with reliability.

Since $\theta^{R*} = q$, The optimal θ^{R*} has to fulfill the equation $l(\theta^{R*}) = v(\theta^{R*})$, i.e.,

$$\frac{\beta}{\delta^2}\theta^{R*} - (1+\theta^{R*})^\gamma = V - \frac{B}{\delta^2}. \tag{24}$$

Since v is strictly concave, l is linear, $v(0) = V + 1 \geq \frac{B}{\delta^2} = l(0)$ and there exists an intersection point of the two curves at θ^{R*}, we have that $v'(\theta^{R*}) < l'(\theta^{R*})$. If γ increases, then v increases and the intersection point θ^{R*} shifts to the right. If we increase β, then l

increases and the intersection point θ^{R*} shifts to the left. A similar logic applies to the other parameters.

Note that the Proposition holds independently of the assumption that the objective function is increasing in worker competence for the parameters δ, β, γ. Because of the assumption $V+1 \geq \frac{B}{\delta^2}$, $s = 0$. f may have a maximum in B. Then, $\theta^{R*} = p$. 1. If we increase γ, then p increases. Second, if we increase β, then p decreases. These two facts follow from the formula of p and the restriction $\theta^{R*} = p \in (0,1)$. θ^{R*} can either be p or q (depending on $p \in B$ or $p > q$). ■

Proposition 4 *Suppose that there is an upper bound on the level of competence compatible with reliability. The manager prefers reliability (with less than full competence) to spot interaction (with full competence) if the optimal level of competence compatible with reliability is high enough.*

Proof of Proposition 4

The proposition simply follows from the fact that M's objective function is monotonically increasing in W's competence. To explicitly derive the threshold level of competence, note that reliability (with competence θ^{R*}) is preferred to spot interaction (with full competence) if

$$f(\theta^{R*}) = -\alpha\theta^{R*} + (1+\theta^{R*})^{\gamma} + V - \frac{1-\delta}{\delta}(\beta\theta^{R*} + B) > -\alpha + b(1) = -\alpha + \beta + B. \quad (25)$$

Plugging in from equation (24) for the optimal θ^{R*} we have, after rearranging,

$$f(\theta^{R*}) = -\alpha\theta^{R*} + \beta\theta^{R*} \underbrace{\left(\frac{1}{\delta^2} - \frac{1-\delta}{\delta}\right)}_{>1} + B \underbrace{\left(\frac{1}{\delta^2} - \frac{1-\delta}{\delta}\right)}_{>1}. \quad (26)$$

Thus, M prefers an agent with less than full competence to a highly competent one, if

$$\theta^{R*} > \frac{-\alpha + \beta + B\left(\frac{-1}{\delta^2} + \frac{1}{\delta}\right)}{-\alpha + \beta\left(\frac{1}{\delta^2} - \frac{1-\delta}{\delta}\right)} \quad (27)$$

and vice versa. Note that if $-\beta < B < \frac{-\alpha+\beta}{\frac{1}{\delta^2} - \frac{1}{\delta}}$, the threshold for θ^{R*} is in $(0,1)$.

■

Proposition A.1 If $-\alpha < \frac{1-\delta}{\delta}\beta$, $f'(\theta) < 0 \; \forall \theta > p := \text{argmax}_{x \in [0,\infty)} f(x)$. Note that $p = \left(\frac{\alpha + \frac{1-\delta}{\delta}\beta}{\gamma}\right)^{\frac{1}{\gamma-1}} - 1$. Then, three cases need to be distinguished: (1) If the function f has a maximum in B, we can use the formula $\theta^{R*} = p = \left(\frac{\alpha + \frac{1-\delta}{\delta}\beta}{\gamma}\right)^{\frac{1}{\gamma-1}} - 1$ in order to find that $\frac{\partial \theta^{R*}}{\partial \delta} > 0$, $\frac{\partial \theta^{R*}}{\partial B} = 0$, $\frac{\partial \theta^{R*}}{\partial \beta} < 0$, $\frac{\partial \theta^{R*}}{\partial V} = 0$, $\frac{\partial \theta^{R*}}{\partial \gamma} > 0$. (2) If $\text{argmax}_{x \in [0,\infty)} f(x) < s$, then $\theta^{R*} = s$ and we obtain $\frac{\partial \theta^{R*}}{\partial \delta} = 0$, $\frac{\partial \theta^{R*}}{\partial B} > 0$, $\frac{\partial \theta^{R*}}{\partial \beta} > 0$, $\frac{\partial \theta^{R*}}{\partial V} < 0$, $\frac{\partial \theta^{R*}}{\partial \gamma} < 0$ with the same reasoning as in the case with $V+1 \geq \frac{B}{\delta^2}$. (3) If $\operatorname{argmax}_{x\in[0,\infty)} f(x) > q$, then $\theta^{R*} = q = 1$ and we get $\frac{\partial\theta^{R*}}{\partial\delta} = 0$, $\frac{\partial\theta^{R*}}{\partial B} = 0$, $\frac{\partial\theta^{R*}}{\partial\beta} = 0$, $\frac{\partial\theta^{R*}}{\partial V} = 0$, $\frac{\partial\theta^{R*}}{\partial\gamma} = 0$.

Proposition 6 *Consider the case of a constant value of reliability. Reliability is feasible if and only if*

$$\theta \leq 1 - \frac{b-\delta^2 v}{\delta b}. \tag{28}$$

Even when reliability is feasible and has positive net value, the manager may prefer spot interaction with a more competent worker. The more productive the manager is (i.e., $(b-\alpha)$ is large) or the tighter the bound on the feasibility of reliability in terms of the allowed competence levels is (i.e., $\left(1-\theta^{R*}\right)$ is large), the more valuable reliability must be to the manager in order to be preferred over spot interaction, where full competence is feasible. The empirical predictions parallel those of the general model.

Proof of Proposition 6

The non-reneging constraints are very similar as before. In a state where the side opportunity of value b actually arises, the worker's non-reneging constraint (NR-W) is[20]

$$w + x + \frac{\delta}{1-\delta}(w+x) \geq w + b + x + \frac{\delta}{1-\delta}(w+\theta b). \tag{29}$$

Simplifying yields

$$x \geq \frac{b}{\delta} - b(1-\theta) = \theta b + b\frac{1-\delta}{\delta}. \tag{30}$$

In other words, the required reliability rewards are the expected value of side opportunities plus some extra amount that decreases with the discount factor and increases with the value of b. Let $\tilde{x}(\theta) = \theta b + b\frac{1-\delta}{\delta}$ denote the minimal rewards that induce reliability from type θ.

The manager's non-reneging constraint is

$$y + v - x - w + \frac{\delta}{1-\delta}\left[E_\theta y + v - x - w\right] \geq y + v - w + \frac{\delta}{1-\delta}\left[E_\theta y - w\right] \tag{31}$$

or

$$x \leq \delta v. \tag{32}$$

By combining the non-reneging constraints, we obtain the proposition.

Denote with $\theta^{R*} = 1 - \frac{b-\delta^2 v}{\delta b}$ the maximal level of competence compatible with reliability. Full competence is only compatible with reliability if the gains to reliability sufficiently outweigh the value of side opportunities, i.e., if $v > b/\delta^2$, which corresponds to the condition in Proposition 2. The results for spot interaction and contractible reliability translate directly to this case. In both instances, it can be easily seen that maximal competence is optimal.

Now consider the optimal choice of M. Because M is relatively productive, it is optimal for M to choose the maximal competence level compatible with reliability, $\theta^{R*} = 1 - \frac{b-\delta^2 v}{\delta b}$.

[20] Note that the non-reneging constraint when no opportunity arises is implied by NR-W.

From this one can show that the parallel result of Proposition 5, namely that: Even when reliability is feasible and has positive net value, the manager may prefer spot interaction with a more competent worker. This result is not simply due to the fact that reliability's absolute cost increases with competence. Instead, it arises because higher competence can be achieved in spot interaction than in reliability.

Specifically, the payoff for M under reliability then is

$$\begin{aligned} U_M^R\left(\theta^{R*}\right) &= E_{\theta^{R*}}y - w\left(\theta^{R*}\right) + v - x\left(\theta^{R*}\right) \\ E_{\theta^{R*}}y - \bar{E}_{\theta^{R*}}y - \theta^{R*}\alpha + \theta^{R*}b + v - \theta^{R*}b - b\frac{1-\delta}{\delta} &= v - b\frac{1-\delta}{\delta} - \theta^{R*}\alpha. \end{aligned} \tag{33}$$

Plugging in $\theta^{R*} = 1 - \frac{b-\delta^2 v}{\delta b} = \delta\frac{v}{b} - \frac{1-\delta}{\delta}$, we get

$$\begin{aligned} U_M^R\left(\theta^{R*}\right) &= v - b\frac{1-\delta}{\delta} - \alpha\delta\frac{v}{b} + \frac{1-\delta}{\delta}\alpha = \\ v\left(1 - \frac{\alpha\delta}{b}\right) - (b-\alpha)\frac{1-\delta}{\delta} &= v\left(1 - \frac{\alpha\delta}{b}\right) - \frac{b-\alpha}{\delta} + b - \alpha. \end{aligned} \tag{34}$$

Compare this payoff to the one obtainable under spot interaction. There, we know that $\theta = 1$ is optimal, leading to a payoff of

$$U_M^S(\theta = 1) = b - \alpha. \tag{35}$$

Therefore,

$$U_M^R\left(\theta^{R*}\right) > U_M^S(\theta = 1) \Leftrightarrow v\left(1 - \frac{\alpha\delta}{b}\right) > \frac{b-\alpha}{\delta} \tag{36}$$

or

$$v > b\frac{b-\alpha}{\delta b - \delta^2\alpha} = \hat{v}(b, \alpha, \delta). \tag{37}$$

We can easily verify that $\hat{v}(b, \alpha, \delta) > b$, i.e., that $\frac{b-\alpha}{\delta b - \delta^2\alpha} > 1$ as $\alpha < 0$.[21]
We can also rewrite the condition that reliability is preferred to spot interaction as

$$v - b\frac{1-\delta}{\delta} - \theta^{R*}\alpha > b - \alpha \tag{38}$$

$$v - b\frac{1-\delta}{\delta} - \theta^{R*}b > b - \theta^{R*}b - \alpha + \theta^{R*}\alpha \tag{39}$$

$$v - x\left(\theta^{R*}\right) > (b-\alpha)\left(1 - \theta^{R*}\right) = \frac{(b-\alpha)\left(b - \delta^2 v\right)}{\delta b} > 0 \tag{40}$$

[21] We can also check that as δ goes to unity, the area where $\theta^L < 1$ is required vanishes and gets "squeezed out" between the $v = b/\delta^2$ and $\hat{v}(b, \bar{\alpha}, \delta)$ schedules. Another way to look at the $\hat{v}(b, \bar{\alpha}, \delta)$ schedule is to ask what happens to competence as we move up the graph and to the right along $\hat{v}(b, \bar{\alpha}, \delta)$. It is easy to check that θ^L decreases, as indicated by the arrows along $\hat{v}(b, \bar{\alpha}, \delta)$ in the Figure. Along $\hat{v}(b, \bar{\alpha}, \delta)$, as b goes to infinity, θ^L approaches $2 - \frac{1}{\delta}$.

where the inequality follows from the premise $b - \delta^2 v > 0$. Thus, reliability needs to have strictly positive net value; merely surpassing zero by some small amount ε is not enough.

Recall that the contractible reliability (static multi-tasking) model prescribes full competence and reliability whenever reliability has positive social value, $v > b$. Figure A.1 describes the different implications that arise for the case of noncontractibility.

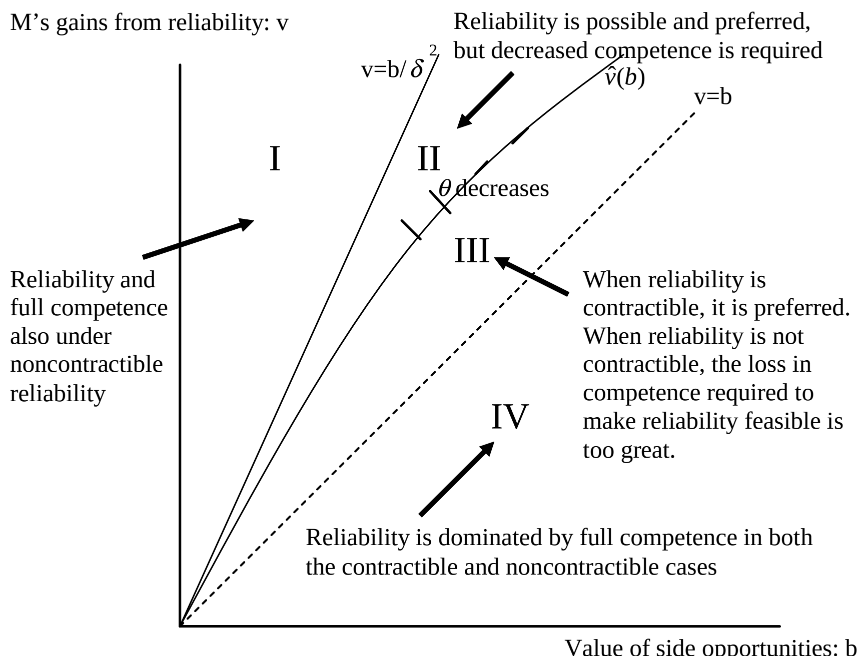

Figure A.1: The tradeoff between competence and reliability and its dependence on M's gains from reliability and W's opportunity cost of reliability. The graph approximately assumes $\delta = 0.75, \alpha = -0.4$, and plots b between 0 and 3.

The graph has four relevant regions. Begin on the top left of the graph. In region I, we have $b - \delta^2 v < 0$, i.e., we are to the left and above of the $v = b/\delta^2$ line. The expression for θ^{R*} tells us that we can obtain full competence and reliability. This is preferred to full competence alone.

As $b - \delta^2 v$ turns positive, we enter region II. As we move down and to the right of the graph (as b increases relative to v), θ^{R*} must become smaller than unity. At some point, the losses in competence required to still obtain reliability become too great, and M is better off not trying to obtain reliability, but focus instead on full competence under spot interaction. We saw above that $\hat{v}(b, \alpha, \delta) = b\frac{b-\alpha}{\delta b - \delta^2 \alpha}$ is the minimum level of reliability gains required to make reliability desired for given other parameter values.

Below $\hat{v}(b, \alpha, \delta)$, in region III of the graph in Figure A.1, full competence is preferred because the optimal choice of competence in order to retain reliability is unattractively low. There are two interesting sub-regions, though. Far below $\hat{v}(b, \alpha, \delta)$, namely, where $v < b$ (in region IV), competence dominates reliability also under contractible reliability. However, between the 45-degree line and $\hat{v}(b, \alpha, \delta)$, if M could contract on reliability, she would be better off doing so. But non-contractible reliability forces her to give up too much competence and thus induces her to forego reliability. This area thus shows the losses due to the need for a self-enforcing reliability contract.

From the above, we know that for M to prefer reliability with competence θ^{R*} to spot interaction with competence $\theta^S = 1$, we must have $v - x\left(\theta^{R*}\right) > (b - \alpha)\left(1 - \theta^{R*}\right) > 0$. Thus, the more productive M is (i.e., $(b - \alpha)$ is large) or the tighter the bound on the feasibility of reliability in terms of the allowed competence levels is (i.e., $\left(1 - \theta^{R*}\right)$ is large), the more valuable reliability must be to M in order to be preferred over spot interaction, where full competence is feasible.

The empirical predictions parallel those of the general model. Specifically, the more productive M is ($-\alpha$ is larger), the more likely he is to implement spot interaction, which is associated with (weakly) higher reliability than any reliability equilibrium. A lower value of side opportunities, b, a higher discount factor, δ, and a higher value of reliability to the manager, v, all increase the maximum level of competence compatible with reliability, θ^{R*}. When reliability is optimal, these comparative statics also hold for the level of competence we expect to observe in equilibrium. The model also has predictions for two interaction terms: The interaction between side opportunities and length of expected interaction reduces competence, while the interaction between the value of reliability and length of expected interaction increases competence. Mathematically, this can be seen from the cross-derivatives of θ^{R*}. Intuitively, the logic behind these two predictions is the following. When the duration of interaction is longer, competence is, ceteris paribus, higher. On the other hand, as side opportunities become greater, the deviation temptation becomes greater, decreasing optimal competence. Thus, with a long duration of interaction, the detrimental effect of an increase in side opportunities is bigger than if competence is already relatively small (which is the case when the duration of interaction is short). The opposite holds for the value of reliability.
■

Proposition 7 *Consider the case of a constant value of reliability. A necessary and sufficient condition for the tradeoff between reliability and competence to arise is that the manager cannot commit to reward the worker for reliability.*

Proof of Proposition 7

For sufficiency, suppose that M's non-reneging constraint, $x \leq \delta v$, remains valid, while instead of W's non-reneging constraint, only W's participation constraint, $x \geq \theta b$, needs to hold. That is, M offers pay for performance, but cannot commit to deliver this promised pay. Combining these two constraints, the upper bound on the level of competence compatible with reliability when M is tempted, θ_L^M, is

$$\theta_L^M = \delta \frac{v}{b} = 1 - \frac{b - \delta v}{b}. \tag{41}$$

In this case, reliability is preferred to spot interaction if

$$v - \theta_L^M \alpha > b - \alpha, \tag{42}$$

or

$$v > b\frac{b-\alpha}{b-\delta\alpha} = \delta\hat{v}\left(b,\alpha,\delta\right). \tag{43}$$

This implies the same comparative statics as for the main case.

For necessity, suppose that only W has a commitment problem while M can commit to pay an "efficiency wage" before W's action. Thus, the two relevant constraints are W's non-reneging constraint, $x \geq \theta b + b\frac{1-\delta}{\delta}$, and the affordability constraint that $v \geq x$. From these two constraints, the upper bound on the level of competence compatible with reliability when W is tempted, θ_L^W, is

$$\theta_L^W = 1 - \frac{b-\delta v}{\delta b}. \tag{44}$$

Thus, reliability is preferred to spot interaction if

$$v - b\frac{1-\delta}{\delta} - \theta_L^W\alpha > b - \alpha, \tag{45}$$

or

$$v > \frac{b}{\delta} \tag{46}$$

Now, if indeed $v = \frac{b}{\delta}$, then $\theta_L^W = 1$. In other words, when reliability is preferred, this also implies that, in fact, full competence is possible. Therefore, no tradeoff between reliability and competence exists for M's optimal choice in this instance. ■

Proposition 8 *Consider the basic signaling game. There exists a separating perfect Bayesian equilibrium that is characterized by the following features:*
1) The market believes that θ_H is independent and that θ_L is reliable. The two types behave accordingly.
2) The reliability rewards x and wages w_H, w_L satisfy the following conditions:
a) High is at least as well off as Low. That is: $w_H + \theta_H b \geq w_L + x$.
b) Low does not want to be independent. That is: $w_L + x > w_H + \theta_L b$.
c) High exactly receives his reservation utility. That is: $w_H + \theta_H b = \bar{u}\left(\theta_H\right)$.
d) Low earns a surplus over his reservation utility. That is: $w_L + x > \bar{u}\left(\theta_L\right)$.
(3) The manager never does better than if the market has symmetric information.

Proof of Proposition 8

To provide a better overview, the proof is divided into several steps.
Optimality of Low's strategy. Suppose that the market has the beliefs as stated, i.e., that θ_L will be reliable and θ_H will be independent. Then, we require two similar but conceptually distinct conditions for Low. First, we require him to be reliable rather than disloyal. By being reliable, W knows that the market will believe him to be θ_L, and, consistent with equilibrium, it believes that it is optimal for M to consequently offer wage w_L and reliability rewards x. By contrast, independence allows the worker to convince the market that he is θ_H. Recall from the text that we assume that even though M knows that the worker is the low type, in order to keep him inside the firm she would need to offer a wage w_H forever.

Low is not really a high type, though. Therefore, in spot interaction, he will expect to earn $\theta_L b$ in side opportunities. In other words, incentive compatibility for Low, IC-L, under which Low prefers reliability to independence holds if and only if

$$w_L + x + \frac{\delta}{1-\delta}(w_L + x) \geq w_H + b + \frac{\delta}{1-\delta}(w_H + \theta_L b). \tag{47}$$

Second, Low must not have an incentive to renege, in addition to "just" mimicing. Using the standard logic, we have for Low's non-reneging constraint, NR-L,

$$w_L + x + \frac{\delta}{1-\delta}(w_L + x) \geq w_L + b + x + \frac{\delta}{1-\delta}(w_H + \theta_L b). \tag{48}$$

Thus, a Low type knows that reneging on reliability (or choosing independence over reliability) has two effects: On the one hand, he is able to secure a higher wage. On the other hand, he loses reliability rewards x. Both weigh more for incompetent guys, so it is a priori unclear which effect dominates. For separation to become possible we need to find values of x and the wages that are consistent with M's optimizing behavior.

To do this, we begin by noting that an immediate implication of NR-L is $w_L + x \geq w_H + \theta_L b$, since $\frac{1-\delta}{\delta} > 0$. This proves part 2b. Moreover, we have the following relationship between Low's two constraints.

Lemma 8.1 *If Low does not renege on reliability, he chooses reliability from the beginning. That is, NR-L implies IC-L.*

Proof of Lemma 8.1

To see this, rewrite NR-L as

$$x \geq (w_H - w_L) + \theta_L b + b\frac{1-\delta}{\delta}.$$

Therefore, $w_L + x > w_H$ is guaranteed, which is a sufficient condition for IC-L to be implied by NR-L. The intuition is clear: If Low does not even have an incentive to renege, he surely does not have an incentive to choose independence from the beginning. ■

Optimality of High's strategy. High needs to prefer independence in this candidate equilibrium. But in fact M can always easily make that constraint hold. She can just never reward reliability by High. In other words, she can just set rewards High gets for reliability very small or negative. If instead IC-H does have to hold, and the rewards to reliability are constrained to be the same for all types, a separating equilibrium may fail to exist. See Proposition 8' below.

Optimality of M's strategy, part 1. M's behavior also needs to obey a non-reneging constraint of the usual form. The slight difference to the constraint in the case where the market knows the worker's competence is that once M has reneged on her promise of rewards for reliability and W is never reliable again, the market will believe that W is θ_H, thus forcing M to pay w_H. Therefore, M's non-reneging constraint, NR-M, is

$$y + v - x - w_L + \frac{\delta}{1-\delta}[Ey + v - x - w_L] \geq y + v - w_L + \frac{\delta}{1-\delta}[Ey - w_H] \tag{49}$$

Rewriting the constraint yields

$$x \leq \delta \left(w_H - w_L\right) + \delta v. \tag{50}$$

Thus, as before, the two non-reneging constraints NR-L and NR-M impose bounds on reliability rewards. The lower bound is due to the W's demands for rewards; the upper bound is due to the lack of full commitment by M.

Utility levels. Intuitively, separation is meaningful if we can induce High and Low to act according to the market's belief (and according to what M found optimal under perfect information) but still retain High's utility advantage. Indeed, we can show parts 2a and 2c of the Proposition in the following

Lemma 8.2 *High earns exactly his reservation utility and is better off than Low. That is,* $w_H + \theta_H b \geq w_L + x$.

Proof of Lemma 8.2

First, we can note that the wage M will pay High will be such that together with his side opportunities, he exactly reaches his reservation utility. There would be no point in paying more (nor less, because High would leave M). Thus, we know that

$$w_H = \bar{u}\left(\theta_H\right) - \theta_H b \tag{51}$$

which is part 2c of Proposition 8.

Second, making use of NR-L, we can substitute for the right side of the condition to be shown in order to have

$$w_H + \theta_H b \geq w_H + \theta_L b + b\frac{1-\delta}{\delta}. \tag{52}$$

Rewriting, we get

$$\left(\theta_H - \theta_L - \frac{1-\delta}{\delta}\right) b > 0 \tag{53}$$

which always holds under our restrictions on parameters. Intuitively, what this condition says is that the types we want to separate must not be "too close," relative to a measure of impatience. ■

Moreover, we can show that Low earns a surplus over his reservation wage (part 2d of Proposition 8).

Lemma 8.3 *Low earns a surplus over his reservation wage:* $w_L + x \geq \bar{u}\left(\theta_L\right).$

Proof of Lemma 8.3

To show this, we combine Low's NR with the insight that High earns exactly his reservation wage. Formally,

$$w_L + x \geq w_H + \theta_L b + b\frac{1-\delta}{\delta} = \bar{u}\left(\theta_H\right) - \theta_H b + \theta_L b + b\frac{1-\delta}{\delta}. \tag{54}$$

This term is greater than $\bar{u}(\theta_L) = \theta_L \bar{y} + (1 - \theta_L)\underline{y}$, if and only if

$$(\theta_H - \theta_L)(\bar{y} - b) + b\frac{1-\delta}{\delta} > 0. \tag{55}$$

The assumed parameter restriction $\bar{y} - b > 0$ is therefore a sufficient condition for the Lemma to hold. ■

Optimality of M's strategy, part 2. Finally, we need to check whether M indeed "wants" θ_L to be reliable and θ_H to be disloyal, i.e., whether W's behavior and the market's beliefs are consistent with optimizing behavior on M's part. Employing the fact that $w_H = \underline{y} + \theta_H \Delta y - \theta_H b$, we know that the low type, when he pretends to be θ_H, actually receives less utility than a true high type, because he only realizes opportunities of expected value $\theta_L b$. The payoffs to M are easily determined as follows:

$$U_M(\theta_L \text{ is reliable}) = v + b\left(\theta_H - \theta_L - \frac{1-\delta}{\delta}\right) - (\theta_H - \theta_L)\Delta y \tag{56}$$

$$U_M(\theta_L \text{ is independent}) = \theta_H b - (\theta_H - \theta_L)\Delta y \tag{57}$$

Thus, reliability is preferred for the low type if and only if

$$v > b\left(\theta_L + \frac{1-\delta}{\delta}\right) = x^P(\theta_L). \tag{58}$$

Similarly,

$$U_M(\theta_H \text{ is independent}) = \theta_H b \tag{59}$$

$$U_M(\theta_H \text{ is reliable}) = v - b\frac{1-\delta}{\delta}. \tag{60}$$

Thus, M wants the high type not to be reliable if and only if

$$v < b\left(\theta_H + \frac{1-\delta}{\delta}\right) = x^P(\theta_H). \tag{61}$$

But these two conditions are precisely what it means for High's reliability not to be worth it under perfect information, but for Low's reliability to be worth it. In other words, if and only if reliability pays under symmetric information, M is also able to find a way to separate High and Low in behavior when the market sends incentives to Low to mirror High. M's ability to do this derives endogenously from the tradeoff between competence and reliability. However, M incurs extra costs in doing this to the extent that Low, by acting independently, expects to be able to receive w_H for an extended period of time.

This completes the proof of Proposition 8. ■

Proposition 8' *Under the conditions of Proposition 8, a separating equilibrium with the properties of the equilibrium in Proposition 8 exists also in the case where the manager is asymmetrically informed about the worker's competence.*

Proof of Proposition 8'

Consider now the case where M does not know W's competence either. Even in this case, M may use the sorting possibility presented to him in form of the competence-reliability tradeoff to her advantage. One might think that separation now becomes more difficult. After all, recall that in the case so far we were able to avoid High envying Low because M could simply give very negative rewards for reliability to High. But here, if M has to offer two packages of compensation – one with and one without reliability rewards – High may envy Low. Intuitively, the two packages must be sufficiently different for separation to be possible, but if the types are too close to each other, this might not be possible. However, this proof reveals that in fact the conditions for Proposition 8 are enough to allow us to obtain separation even in this case. Similar conditions as in Proposition 8 need to hold for Low's and M's behavior. In addition, now High needs to prefer independence in this candidate equilibrium. For High, there is no non-reneging constraint, because there is no contract to renege on. His incentive-compatibility constraint, IC-H, is

$$w_H + b + \frac{\delta}{1-\delta}\left(w_H + \theta_H b\right) \geq w_L + x + \frac{\delta}{1-\delta}\left(w_L + x\right) \tag{62}$$

or

$$x \leq \left(w_H - w_L\right) + b\left(1-\delta\right) + \delta\theta_H b. \tag{63}$$

Intuitively, if reliability rewards are too attractive, even high types choose reliability over spot interaction. A necessary condition for both NR-L and IC-H to hold is that

$$\delta\theta_H - \theta_L \geq \delta + \frac{1}{\delta} - 2 \geq 0 \tag{64}$$

This condition can be written as

$$\theta_H > \left(\frac{1}{\delta} - 1 + \theta_L\right) - \left(1 - \delta - \theta_H + \delta\theta_H\right). \tag{65}$$

The term $\left(1 - \delta - \theta_H + \delta\theta_H\right)$ is always positive. Therefore, if $\theta_H > \left(\frac{1}{\delta} - 1 + \theta_L\right)$, the premise of Proposition 8, separation is also possible in the case where M does not know the worker's competence. ■